Marking Scheme for Final Exam in SLR

1. Import the libraries and read the dataset (tab, csv, xls, txt, inbuilt dataset)

2. Summarize important observations from the data set (5 Marks)

3. Check for defects in the data. Perform necessary actions to 'fix' these defects(5 Marks)
    - Some pointers to help, but don't be limited by these
        - missing/null values
        - variables have outliers?
        - Is the data normally distributed, etc
        -
4. Summarize relationships among variables(10 Marks)(10 marks)
    - You can plot correlation plots, pair plots etc

5. Do the Data preprocessing including the following :(10 marks)
        - Separate features and target
        - Encode categorical variables
        - Scale the variable if needed
        - Split the data into train and test (70: 30)
        -
6. Fit a base model. Please write your key observations (10 marks)

7. How do you improve the accuracy of the model? Write clearly the changes that you will make before re-fitting the model. Fit the final model. (20 marks)

8. Summarize as follows below. (10 marks)
        - Summarize the overall fit of the model and list down the measures to prove that it is a good model
        - Write down a business interpretation/explanation of the model – which variables are affecting the target the most and explain the relationship. Feel free to use charts or graphs to explain.
        - What are the key risks to your results and interpretation

****